The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations

Ryde, Ulf

Published in:
Journal of Computer-Aided Molecular Design

DOI:
10.1007/BF00402823

1996

Document Version:
Peer reviewed version (aka post-print)

The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum chemical and molecular mechanical calculations

by

Ulf Ryde

Department of Theoretical Chemistry

University of Lund

Chemical Centre

P.O.B 124

S-221 00 Lund

Sweden

Tel: 46-46-2224502

Fax: 46-46-2224543

Email: teoulf@garm.teokem.lu.se

Key words: geometry optimisation, five-coordination, reaction mechanism, geometry imposed by enzyme, protein strain

Abstract

The coordination number of the catalytic zinc ion in alcohol dehydrogenase has been studied by integrated ab initio quantum chemical and molecular mechanical geometry optimisations involving the whole enzyme. A four-coordinate active-site zinc ion is 100-200 kJ/mole more stable than a five-coordinate one, depending on the ligands. The only stable binding site for a fifth ligand at the zinc ion is opposite to the normal substrate site, in a small cavity buried behind the zinc ion. The zinc coordination sphere has to be strongly distorted to accommodate a ligand in this site, and the ligand makes awkward contacts with surrounding atoms. Thus, the results give no support to proposals attributing an important role to five-coordinate zinc complexes in the catalytic mechanism of alcohol dehydrogenase.

The present approach makes it possible also to quantify the strain induced by the enzyme onto the zinc ion and its ligands; it amounts to 42-87 kJ/mole for four-coordinate active-site zinc ion complexes and 131-172 kJ/mole for five-coordinate ones. The four-coordinate structure with a water molecule bound to the zinc ion is about 20 kJ/mole less strained than the corresponding structure with a hydroxide ion, indicating that the enzyme does not speed up the reaction by forcing the zinc coordination sphere into a structure similar to the reaction intermediates.

Alcohol dehydrogenase (EC 1.1.1.1) catalyses the reversible oxidation of primary and secondary alcohols using NAD^+ as coenzyme [1-3]. The active site of the enzyme contains a zinc ion that is essential for catalysis. Crystallographic studies [3-5] have shown that this zinc ion is bound by the enzyme through two cysteine and one histidine residue. In free enzyme, the catalytic zinc ion appears to be tetrahedrally coordinated with one water molecule (or hydroxide ion, depending on pH) as the fourth first-sphere ligand.

The most widely accepted reaction mechanism for alcohol dehydrogenase [1] consists of (**Scheme 1**): (1) binding of NAD^+; (2) binding of the alcohol substrate by replacing the zinc-bound water molecule; (3) deprotonation of the alcohol; (4) hydride transfer from the alkoxide ion to NAD^+, yielding NADH and a zinc-bound aldehyde; (5) release of the aldehyde by a replacing water molecule; (6) dissociation of NADH. In applicable steps (7-9), the zinc-bound water molecule equilibrates with a zinc-bound hydroxyl ion.

According to this mechanism, the active-site zinc ion remains four-coordinate in all significant catalytic steps. Alternative proposals have been put forward, however, according to which five-coordinate intermediates play an essential role during catalysis [6-13]. For example, Eklund & Brändén [3] and Merz [12] have suggested that the alcohol substrate is deprotonated by an internal proton transfer within a five-coordinate complex of the alcohol and a hydroxyl ion. Other authors have proposed that this complex [10-11,14] or the resulting five-coordinate alkoxide-water complex [6,13], is the intermediate undergoing the catalytic hydride transfer. The ideas have also been advanced that five-coordinate zinc complexes with an alkoxide ion and hydroxyl [6] or hydronium ion [9] are involved in the catalytic reaction mechanism.

Crystallographic studies of the enzyme and its binary or ternary complexes with coenzyme and different substrates have shown that the catalytic zinc ion as a rule exhibits four-coordination [2-5,15,16], and several spectroscopic investigations have provided evidence in the same direction [17-24]. On the other hand, there is strong crystallographic and spectroscopic evidence showing that binding to zinc of certain bidentate inhibitors is five-coordinate [25-26]. Furthermore, spectroscopic studies of metal-substituted alcohol dehydrogenase have indicated that some binary and ternary complexes may be five-coordinate

[8,9,27-34]. The kinetic evidence is also scattered and has been taken to favour four-coordination [1,35], as well as five-coordination [6,8,10-11,13], of zinc in the catalytically productive ternary-complexes.

Recently, Ryde published an extensive series of quantum chemical calculations on models of the active site of alcohol dehydrogenase with a varying number of different non-protein ligands [36]. These calculations indicated that four-coordinate structures were about 20 kJ/mole more stable than five-coordinate ones. Furthermore, no stable five-coordinate complexes with a negative total charge could be obtained. Thus, in vacuum a zinc ion with ligands similar to those found in the enzyme prefers four-coordination over five-coordination. The significance of these results for the reaction of alcohol dehydrogenase is less clear, however, even if the fact that the geometries of the calculated structures are similar to the crystallographic structure of the active site may be taken to indicate that the calculations are relevant also to the conditions in the enzyme.

These quantum chemical calculations were later used to construct a force-field parameterisation of the zinc ion tailored for the active site of alcohol dehydrogenase and this parameterisation was utilised in molecular dynamics simulations and molecular mechanical energy minimisations of the enzyme [37]. They provided a detailed picture of the dynamics of the zinc ligands and indicated that a four-coordinate active-site zinc ion is about 40 kJ/mole more stable than a five-coordinate one in the enzyme.

Although appropriate for molecular dynamics simulations, such a parametrisation necessarily provides a rather crude description of some parts of the potential surface around the zinc ion. Much more detailed structural information can be obtained if the active site is treated quantum chemically. Methods have been developed in which the course of a quantum chemical geometry optimisation is influenced by classical forces exerted by the environment [38-42]. By such approaches it is possible to accurately investigate the influence of the enzyme on the geometry of molecules in the active site. In this paper, such a method is further developed and used to obtain accurate information on the optimal structure and on the energetics of the coordination of the active-site zinc ion of alcohol dehydrogenase.

Methods

Geometry optimisations with a combined quantum chemical and classical method

Here, a general method for combined quantum chemical and molecular mechanical geometry optimisations is described. The approach is similar to the one used in the program QUEST by Singh & Kollman [38].

The total system (enzyme + solvent) is divided into four subsystems:
system 1: a small central system to be described quantum mechanically (the quantum system);
system 2: all atoms of amino acids within radius r_1 of any atom in system 1;
system 3: all atoms of amino acids within radius r_2 of any atom in system 2;
system 4: the rest of the total system.

During the geometry optimisation, system 1 is optimised using the sum of the quantum chemical forces within system 1 and molecular mechanical forces from system 2 onto system 1. All electrostatic interactions between system 1 and system 2-4 are included in the quantum chemical calculations, so that the quantum system is polarised by the charges in system 2-4. Therefore, no molecular mechanical forces due to the electrostatics are calculated. The geometry of system 2 is optimised molecular mechanically once in each iteration of the optimisation of system 1. In this way the relative speed of the molecular mechanical optimisation is exploited (the molecular mechanical minimisation takes about the same time as one quantum chemical wave function or force evaluation), and the number of iterations in the optimisation procedure is kept to a minimum. The coordinates of the atoms in system 3 and 4 are kept fixed.

The flow of the geometry optimisation is shown in **Scheme 2**. First, the quantum chemical forces within system 1 and the classical forces from system 2 onto the atoms in system 1 are calculated (steps 1-2). Second, the sum of these forces is used for relaxation of the geometry of system 1 (steps 3-5). Third, after inclusion of Mulliken charges of system 1 into the classical representation, system 2 is geometry optimised by classical molecular mechanics, keeping the atoms in system 1, 3 and 4 fixed (steps 6-8). Finally, the sum of the quantum chemical and classical energy of system 1 and 2 is calculated (and also the wave

function and Mulliken charges of system 1; steps 9-12). If the change in energy and geometry is below specified thresholds the optimisation is stopped, otherwise a new optimisation cycle is initiated.

Optionally, the coordinates of system 2 can be kept fixed (i.e. the protein is not allowed to relax in effect of the change of geometry of system 1). If so, no Mulliken analysis or molecular mechanics optimisation is performed (steps 6-8, and 10 in Scheme 2).

In the quantum chemical computations (steps 1, 9, and 10 in Scheme 2), system 1 is described by a wave function, while system 2 and 3 are represented by partial charges, one for each atom, and system 4 by integer charges, i.e. one charge for each charged amino acid, located at the position of the NZ, CZ, CG, CD, SG, CE1, ZN and both P atoms of Lys, Arg, Asp, Glu, Cys^-, His^+, ZN and NADH, respectively. The integer charges are damped by a dielectric constant ε=4.0, while for the rest of the system, ε=1.0. All the charges are treated as atoms with a nuclear charge but no basis functions and their interactions (with themselves and with the system 1 nuclei and electrons) are included in the one-electron Hamiltonian of the quantum system. The quantum chemical forces onto the point charges of system 2-4 are discarded.

In the classical energy and force evaluations (steps 2 and 11 in Scheme 2), only system 1 and 2 are included and electrostatic interactions are ignored since they are already treated quantum chemically. Finally, in the molecular mechanical geometry optimisation of system 2 (step 7 in Scheme 2), system 1-3 are included in all-atom representation, using charges obtained from a quantum chemical Mulliken analysis for system 1 and standard partial charges for system 2 and 3, while system 4 is represented by damped integer point charges.

Special action has to be taken when there is a chemical bond between one atom X in system 1 and an atom C in system 2. In the quantum chemical calculations, C is replaced by a hydrogen atom H. The coordinates of this atom x_H is determined from x_X and x_C according to Eqn. (1)

$$x_H = x_X + (x_C - x_X)\frac{HX_0}{CX_0},$$

(1)

i.e. it is ensured that the H-X bond length differs as much from the optimal H-X bond length computed quantum chemically with the same basis sets (HX_0), as the C-X bond length differs from the equilibrium length of the C-X bond in the force-field library (CX_0). Conversely, the position of C is obtained from x_X and x_H by the inverse of Eqn. 1. When system 2 is allowed to relax, the position of hydrogen atoms bound to C (HC) are determined by the molecular mechanical minimisation. If system 2 is kept fixed, the C-HC distances and the angles around C that does not involve X are kept fixed.

The partial charges of the HC hydrogen atoms and the other heavy atom bond to C are set to 0 in the quantum chemical computations (steps 1, 9, and 10 in Scheme 2) and the charges of the rest of the atoms in the amino acid are uniformly scaled so that the total residue charge vanishes (a change of less than 0.006 e / atom). The charge on the C junction atoms in the molecular mechanical energy minimisation of system 2 (step 7 in Scheme 2) is adjusted so that the total charge of system 1-4 is not changed. This means that to the Mulliken charge of each C atom is added the difference between the sum of the charges of the system 1 atoms in the standard classical charge assignment and in the Mulliken analysis, divided by the number of junctions. In this way the total charge of the enzyme is conserved, while charge transfer between the zinc ion and its ligands is allowed and the charge on the C atom is changed from that typical for hydrogen atoms to that typical for a carbon atom.

The total energy of the optimised systems, E_{tot} is given by

$$E_{tot} = E_q + E_c + E_{pol}$$

(2)

Here, E_q is the standard quantum chemical energy of system 1, including the interactions of the classical charges of system 2-4 with themselves and with the nuclei and electrons of system 1 (evaluated in step 9 in Scheme 2). E_{pol} is the energy of the polarization of system 2 by the quantum system, given by [43]

$$E_{pol} = -\frac{1}{2}\sum_{i}^{system\,2} \alpha_i \left|\overline{F_i}\right|^2 ,$$

(3)

where F_i is the quantum chemical electrostatic field of system 1 at the position of atom i in system 2, and α_i is the polarisability of this atom, adapted from Merck Molecular Force Field [44]. E_{pol} is calculated only after the geometry optimisation. The classical energy (computed in step 9 in Scheme 2), finally,

$$E_c = E_{c2} - E_{c1},$$

(4)

is defined as the difference of the molecular mechanical potential energy in two separate calculations (both with all charges = 0), one with system 1-2 and C junction atoms (E_{c2}), and one with only system 1 and H junction atoms (E_{c1}). Similarly, the classical forces (calculated in step 2 in Scheme 2) are the difference of the forces between these two systems. In this way, the classical energy and forces within system 1 are cancelled out; E_c contains only interactions (van der Waals, bond stretching, angle bending, dihedral torsions, etc.; Eqn. 5 below) that involve at least one atom in system 2, and only those forces that are due to interactions involving at least one atom in system 1 and one atom in system 2 are retained, see **Figure 1**. Furthermore, the total energy and forces (classical + quantum chemical) refer to a system with C junction atoms since the quantum chemical calculations involve system 1 with H junction atoms and the terms in the classical forces involving junction atoms are differences between a C junction and a H junction.

In the present implementation (the program COMQUM) the semi-direct program package TURBOMOLE [45] has been combined with the molecular dynamics program MUMOD [37,46]. The interface consists of four small procedures performing steps 3, 5, 6, 8, and 12 in Scheme

2, a program constructing all input files, and a shell script driving the geometry optimisation. No changes have been made to the code of TURBOMOLE or MUMOD.

The energy function of MUMOD is given in Eqn. 5 (i.e. this is E_{c1} and E_{c2} in Eqn. (4), with $q_i=q_j=0$, and also the energy optimised in step 7 in Scheme 2):

$$E = \sum_{bonds} A_i (r_i - r_{i_0})^2 + \sum_{angles} B_i (\alpha_i - \alpha_{i_0})^2 + \sum_{dihedrals} \sum_{j=1}^{3} C_{ij} \left(\cos(j\phi_i) + 1\right) + \sum_{non-bonded\ \_i<j} \left(\frac{D_{ij}}{r_{ij}^{6}} + \frac{E_{ij}}{r_{ij}^{12}} + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}}\right) \quad (5)$$

The first three terms represent the energies of bond stretching, angle bending and dihedral torsions, where r_i, α_i and ϕ_i are the actual bond lengths, angles and dihedral angles and r_{i0} and α_{i0} are the corresponding equilibrium values. The fourth term represents the non-bonded interactions, consisting of a Lennard-Jones 6-12 term and a Coulomb term, where r_{ij} is the distance between atom i and j. 1,4 but not 1,3-interactions are included in the non-bonded potential. The force field does not contain any specific terms for hydrogen bonds or improper dihedral angles. No cut-off of the non-bonded forces was applied. The threshold in the molecular mechanical optimisations (step 7 in Scheme 2) was 10^{-4} kJ/mole/pm for the norm of the gradients and in the whole geometry optimisation 10^{-4} Hartree (0.26 kJ/mole) and 10^{-2} Bohr (0.53 pm) for the change in energy and the Cartesian coordinates, respectively. All optimisations with system 2 free to relax were concluded by an optimisation with system 2 fixed, using tighter thresholds (10^{-6} Hartree and 10^{-3} Bohr).

Several starting geometries were tested in order to minimise the risk of being trapped in local minima. The geometries were taken from classical molecular dynamics and molecular mechanics simulations of alcohol dehydrogenase involving a zinc parametrisation tailored for the active site of the enzyme [37] and were selected to include different geometries of the zinc coordination sphere (e.g. different axial ligands in the five-coordinate complexes). The quantum chemical computations were performed at the Hartree-Fock level with analytical gradients, using basis sets of double-ζ quality for all atoms (H: (31); C, N, O: (5111/31); S, P:

(521111/4111); Zn: (62111111/51111/311)) [47,48]. All calculations were performed on IBM RISC RS/6000 workstations.

The enzyme

Throughout, the coordinates of horse liver alcohol dehydrogenase in complex with NADH and dimethylsulfoxide at 180 pm resolution (R-factor=0.172) [5] were used. This is at present the most accurate structure of alcohol dehydrogenase. The enzyme is in the closed conformation which is the catalytically interesting conformation (all coenzyme-containing complexes in Scheme 1 are generally considered to be in this conformation [1-4]) and also the one to which all reports of a five-coordinate zinc ion refer. The charge assignment and the positioning of hydrogen atoms and water molecules were performed as described before [37]. System 1 consisted of $Zn(SH)_2$(imidazole)$(OH)_{0-1}(H_2O)_{1-2}$ (16-20 atoms) with junctions at the CB atom of Cys46, Cys174 and His67 (from subunit A of the enzyme). In system 2, all amino acids within 300 pm from any atom in system 1 were included, viz. Ser48, Asp49, Gly66, Glu68, Phe93, Phe140, Leu141, Gly173, Gly175, Ile318, Arg369, H2O158, the nicotinamide moiety of NADH and the rest of Cys46, Cys174 and His67 (totally 206 atoms). System 3 was composed of all atoms of residues within 300 pm of any atom in system 2, viz. amino acids number 43-45, 47, 50-53, 57, 59, 63-64, 69, 90, 92, 94-95, 109-110, 115, 116, 139, 142, 146, 170-172, 176, 178-179, 202-203, 292, 294, 317, 319-321, 345-348, 359, 368, 370, crystal water number 5, 8, 21, 35, 55, 58-59, 159-161, 167, 172, the N-ribose and the pyrophosphate moiety of NADH and amino acids 309 and 310 from the other subunit of the protein, in total, 837 atoms. Finally, system 4 comprises 176 integer charges, leading to a total charge of +4.

Results

Performance of the method

The performance of the method was tested by optimising nine hydrogen bonded systems quantum chemically, molecular mechanically and with COMQUM with system 2 free or fixed. The results in **Table 1** and **Figure 2** show that the method performs excellently. The average root-mean-squared deviation between the quantum chemical structure and the COMQUM structures is only 3.5 and 4.5 pm, with system 2 fixed and free, respectively. The error in the hydrogen bond distance is even less, 1.8 and 1.2 pm, respectively, with no systematic error. The slightly better total performance of the calculations with a fixed system 2 is only due to that the quantum chemical structure was used as the starting structure. If worse starting structures had been used, the calculation with system 2 free would have been better.

Alcohol dehydrogenase

Combined quantum chemical and classical geometry optimisations were performed on the active site of alcohol dehydrogenase with five sets of non-protein ligands. The results of these calculations are collected and compared with vacuum optimisations in **Table 2 and 3**. For all complexes, two optimisations were performed: one with a fixed enzyme, followed by one were system 2 was allowed to relax.

Figure 3a shows the optimised structure of the active-site zinc ion with a four-coordinate water ligand. The water molecule is bound to the zinc ion at the bottom of the substrate cleft and is hydrogen bonded to OG of Ser48. The structure is rather similar to the one obtained in vacuum; the bond lengths, bond angles and dihedral angles differ by only 1.0 pm, 4.0° and 16.6°, respectively. The most significant differences are the bond lengths and angles around the zinc ion and the dihedral angles of the water and sulphide hydrogens. In Table 3 the bond lengths and angles of this structure are compared to those obtained in vacuum, and also to those found in the crystallographic structure of the enzyme. The greatest

differences to the latter structure are found in the Zn-ligand bond lengths. These can be attributed to three sources. Firstly, the calculations refer to a temperature of 0 K, while the crystal structure was determined at 277 K. A correction of the temperature would increase all the bond lengths. Secondly, the crystal structure has dimethylsulfoxide as the fourth zinc ligand, the computations a water molecule. According to ab initio vacuum geometry optimisations [37], dimethylsulfoxide binds about 14 pm tighter to the zinc ion than water, leading also to longer Zn-S distances. Thirdly, more extended basis sets in the quantum chemical computations would give longer Zn-O and Zn-N bonds and shorter Zn-S bonds [38].

In the structure in Figure 3a, as well as in all other calculated structures, the thiolate of Cys174 is closer to the zinc ion than the one of Cys46 (228 compared to 235 pm). In the crystal structure, the state of affairs is opposite (231 and 223 pm). This discrepancy can probably be attributed to the uncertainty in the crystal coordinates; In the recent crystallographic structures of the free form of liver alcohol dehydrogenase and of the copper substituted enzyme with NADH and dimethylsulfoxide [49,50], the trend is the same as in the calculations.

If the atoms of system 2 are kept fixed, the similarity is slightly weaker to the vacuum calculation (1.5 pm, 4.5° and 13.6° for bond lengths, bond angles and dihedral angles, respectively) and the enzyme (5.7 pm, 4.9° and 6.4°).

Figure 3b show the structure of a four-coordinate catalytic zinc ion with one hydroxide ligand. The hydroxide ion is located in the substrate site and is hydrogen bonded to HG of Ser48.

With two water ligands, two qualitatively different structures of the active site were studied: one four-coordinate with the second water molecule in the second coordination sphere of the zinc ion, and one five-coordinate complex. In the four-coordinate structure, shown in Figure 3c, the second-sphere water molecule is hydrogen bonded to the first sphere water molecule (by its oxygen). It does not make any further hydrogen bonds since no appropriate acceptors are available in the hydrophobic substrate binding site. The first-sphere water molecule makes the normal hydrogen bond to OG of Ser48.

The five-coordinate structure, shown in Figure 3d is strongly strained, with 131 kJ/mole higher energy than in vacuum. One water molecule occupies the normal substrate site at the bottom of the substrate cleft, making a hydrogen bond to OG of Ser48. The other water molecule occupies a site opposite to the first water molecule, in a small cavity behind the zinc ion (termed the *alternative* site below). It interacts weakly with either OD of Asp49 (relaxed enzyme) or a crystal water molecule (H-O distance 216 or 242 pm) and makes close contacts with Cys46 and Glu68. In order to accommodate the water molecule in this site, the imidazole group of His67 and the crystal water molecule have to move almost 100 pm. This redirects the lone pair of NE2 in His67 away from the zinc ion, weakening this bond. As is shown in Table 2, the geometry around the zinc ion is highly distorted compared to the vacuum structure; the S-Zn-S has decreased by 55° while the two S-Zn-N angles have increased by about 30°. The average deviation of bond lengths, angles and dihedrals is 1.7 pm, 8.6° and 39.0°, respectively. Altogether, the five-coordinate structure is rather awkward and is 128 (95 with a fixed protein) kJ/mole less stable than the four-coordinate structure with one water molecule in the second coordination sphere of the zinc ion.

An optimisation of a four-coordinate structure with a water ligand in the alternative zinc site was also performed. This structure is 76.2 kJ/mole less stable than the structure with the water molecule in the substrate site, indicating that the alternative water site is highly unfavourable also in four-coordinate structures.

With one water molecule and one hydroxide ion as zinc ligands, again two different structures could be obtained, one five-coordinate and one four-coordinate with the water molecule in the second coordination sphere. In the latter structure, shown in Figure 3e, the hydroxide ion is located in the substrate site, hydrogen bonded to HG of Ser48 and to a hydrogen of the water molecule. A four-coordinate structure with the water molecule in the second coordination sphere of the zinc ion, *behind* the zinc ion (i.e. opposite to the hydroxide ion) was also obtained. Yet, this structure was 157 kJ/mole less stable than the other structure, and is therefore probably of minor significance.

The five-coordinate structure with one hydroxide and one water ligand, shown in Figure 3f, is again very strained, 200 kJ/mole less stable than the four-coordinate structure,

and 173 kJ/mole more strained than in vacuum. The water molecule occupies the substrate site and forms the normal hydrogen bond to OG of Ser48. The hydroxide ion occupies the alternative site and makes no favourable interactions with the protein. The imidazole ring of His67 is strongly tilted and makes an angle of 65° to the zinc ion. Further, the structure is labile; when the protein was allowed to relax, the structure reorganised to a four-coordinate structure with the hydroxide ion in the substrate site and the water molecule in the second coordination sphere. No five-coordinate structures with the hydroxide ion in the normal substrate site could be obtained.

Discussion

The approach

In this paper, a method to integrate quantum mechanical and molecular mechanical geometry optimisations is presented. In the present implementation, the quantum chemical program package TURBOMOLE [45] has been combined with the molecular dynamics simulation program MUMOD [46]. Yet, the method is general and applicable to any combination of quantum chemical and classical mechanical programs. Furthermore, no changes in the code of any of the programs are necessary. The only requirements are that: 1. the quantum chemical program must accept a large number (of the order of 1000) of point charges, 2. the quantum chemical forces must be available on a file before the relaxation is performed, and 3. molecular mechanical forces must be written to a file.

The approach is similar to the one used in the program QUEST [38] and developments thereof [39-42], which have been thoroughly tested and shown to perform well in several enzymic systems. The treatment of the junctions is more consistent, however; In QUEST, there is no correspondence between the H and the C atom of the junction. The H atom is optimised only by the quantum chemical forces, while the C atom is kept fixed. Furthermore, the atoms in system 1 are influenced by forces both from the H atom and from the C atom. In the present approach, the position of C is determined from the position of H (through Eqn. 1) and all forces are corrected to involve only the C junction atom and not the H atom. Thus the positions of the C atoms are also optimised.

Furthermore, the current program allows the surrounding enzyme (system 2) to relax in effect of geometry changes in system 1. This is effected, not by an integrated relaxation protocol involving both the classical and the quantum chemical systems as suggested by Singh & Kollman [38], but by a full molecular mechanical energy minimisation of system 2 in each iteration of the geometry optimisation of system 1. Such an approach is advantageous since the evaluation of the energy and forces in the quantum chemical system is much more expensive than in the classical system; each molecular mechanical minimisation takes about the same time as one quantum chemical wave function or force evaluation. Furthermore, an

integrated protocol would have taken more iterations to converge (of the order of the number of atoms in system 1 and 2) than the present approach (of the order of the number of atoms in system 1).

The test cases in Table 1 and Figure 2 show that with these two improvements the method performs excellently. Yet, in the calculations with alcohol dehydrogenase, the effect of relaxing the enzyme is unexpectedly small. While the classical energy of system 2 changes by 598-635 kJ/mole compared to the optimisation with a fixed enzyme, the quantum chemical energy of system 1 changes by only 5-18 kJ/mole. Considering that calculations with a relaxing enzyme are about three times as expensive as those with fixed enzyme even if less tight convergence criteria are applied, relaxation the protein may not seem fully worthwhile. Perhaps a better approach would be to let the protein relax only a few times during the geometry optimisation (in order to ensure correct location of the polar hydrogen atoms).

The present calculations involve integer charges in system 4. These were used (instead of partial charges) for three reasons. Firstly, partial charges of amino acids vary appreciably between different force fields (by as much as 0.7 e; sometimes even by the sign, e.g. compare the charges of Bellido & Rullman [51] with those in MUMOD [46] or AMBER [52]); integer charges are much more well defined. Secondly, the integer charges give a smoothed cut-off for the electrostatic interactions; interactions with dipoles are cut at one radius, while a much larger radius is used for the charge interactions. Thirdly, quantum chemical programs are often not intended to be used with a large number of point charges. Point charges are treated as atoms without basis functions and if the number of atoms is increased, many other vectors in the program also increase (unnecessarily). If the number of atoms is increased above about 1500, severe memory problems are accounted in Turbomole as well as in many other quantum chemical program packages, thus giving a software limit on the number of point charges that may be used.

The effect of increasing the number of point charges in the quantum chemical calculations by about 800 (0.3 nm increase in the radius) was tested on $Zn(HS)_2$(imidazole)$(H_2O)_2$ and the corresponding four-coordinate structure (with the enzyme fixed). This altered the geometry about 10 pm (root-mean-squared) and increased the relative stability of the

four-coordinate structure by 5 kJ/mole. These effects are smaller than if the partial charges of another force field were used.

The coordination number of the catalytic zinc ion

Eklund et al. have argued on the basis of model building experiments that there is no room for a fifth ligand at the catalytic zinc ion of alcohol dehydrogenase [15]. The present results show that there actually is a fifth binding site at the zinc ion. This site is opposite to the normal substrate site, buried behind the zinc ion in a small cavity delineated by the zinc ion, Cys46, Ser48, Asp49, His67, Glu68, Cys174, Arg369 and a crystal water molecule. Only a molecule of the size of water or smaller may occupy this site [37]. A molecule in this site may make weak hydrogen bonds to the carboxylate group of Asp49 or a crystal water molecule. The cavity is so small that even a water molecule makes awkward contacts with the surrounding atoms (especially with Glu68) and the zinc coordination sphere becomes strongly distorted. This is illustrated by the fact that a four-coordinate water molecule has 106 kJ/mole lower binding energy in the alternative site than in the substrate site.

A five-coordinate complex with a water molecule in the substrate site and a hydroxide ion in the alternative site could also be obtained if the protein was kept fixed. This is a bit unexpected, since five-coordinate complexes with a hydroxide ion are unstable in vacuum [36]. Presumably, the complex is stabilised by the hydrogen bond between the water molecule and Ser48. The structure is probably of minor significance, however, since it could not be obtained when the protein was allowed to relax.

All five-coordinate complexes are severely strained, much more than the corresponding four-coordinate structures (c.f. Tables 2 and 3). The structure with two water molecules is 95-128 kJ/mole less stable than the four-coordinate structure, while the one with a hydroxide ion is 200 kJ/mole less stable. Apparently, four-coordinate zinc structures are favoured not only by the chemical properties of protein zinc ligands (by about 20 kJ/mole [36]) but also by the folding of the enzyme at the active site.

This large difference in the stability of four- and five-coordinate structures in the active site of alcohol dehydrogenase indicates that five-coordinate complexes should be very unstable; Provided that entropic effects do not differ significantly between the two types of zinc complexes (in vacuum, entropy favours five-coordinate complexes by less than 7 kJ/mole [36]), the equilibrium constant for the decay of a five-coordinate zinc complex to a four-coordinate one should be more than $3 \cdot 10^{16}$! This argues strongly against any observation of a five-coordinate catalytic zinc ion in alcohol dehydrogenase and renders improbable all mechanistic proposals involving a five-coordinate zinc ion.

Spectroscopic studies of cadmium, copper, or cobalt substituted alcohol dehydrogenase have in several instances been taken to provide evidence for a five-coordinated catalytic metal site [8,9,27-34]. The present results indicate that, to the extent that five-coordinate complexes do form with metal-substituted enzyme, this probably reflects the disparity in coordination preferences of different metal ions and cannot be taken to suggest that the catalytic metal site in native enzyme is five-coordinate. A similar conclusion has recently been drawn from a comparison between crystallographic and spectroscopic results on the coordination chemistry of the catalytic metal ion in carbonic anhydrase, another zinc enzyme [53].

The strain induced by the enzyme onto the active site

It is widely supposed that the conformation taken by the substrate in the active site of an enzyme is determined mainly by the enzyme [54], i.e. that the enzyme forces the substrate into a conformation appropriate for catalysis. The comparison of a structure optimised quantum chemically in vacuum and with the combined method provides an estimate of the change in geometry and energy when it is inserted into the enzyme. Thus, $?E_{QC1}$ in the third column of Table 2 provides an estimate of the strain forced by the enzyme onto the zinc coordination sphere in alcohol dehydrogenase. This strain amounts to 42-87 kJ/mole for four-coordinate structures and 131-173 kJ/mole for five-coordinate structures, reflecting that the enzyme strongly favours four-coordination.

It is noteworthy that a four-coordinate water ligand is 16-20 kJ/mole less strained than a four-coordinate hydroxide ion. According to the mechanism in Scheme 1, the alcohol substrate must be deprotonated before the hydrogen transfer. Furthermore, structures of zinc complexes with water and alcohols are very similar, as is complexes with hydroxide and alkoxide ions, while these two types of complexes are mutually rather different, especially in the angles subtended at the zinc ion (c.f. Table 3 and ref. no. 35). Together, this seems to indicate that in alcohol dehydrogenase, the enzyme forces the active site into a conformation similar to the reactants and not to the intermediates, i.e. the strain introduced by the enzyme onto the active site disfavours the reaction. The explanation of this apparent maladjustment is probably that the enzyme also has to disfavour zinc-hydroxide complexes which according to the mechanism in Scheme 1 represent dead-end complexes.

Acknowledgements

This investigation has been supported by a grant from the Swedish Natural Science Research Council (NFR) and by IBM Sweden under a joint study contract. The protein coordinates were courteously provided by Eila S. Cedergren-Zeppezauer. The author is indebted to her, Björn O. Roos and Gösta Pettersson for fruitful discussions.

References

1. Pettersson, G. CRC Crit. Rev. Biochem., 21 (1987) 349.
2. Brändén, C.I., Jörnvall, H., Eklund, H and Furugren, B., in the Enzymes, vol. 11A Boyer, P.D. ed., Academic Press, London, 1975, p. 103.
3. Eklund, H and Brändén, in Zinc enzymes Spiro, T.G. ed., John Wiley & Sons, New York, 1983, p. 124.
4. Eklund, H., Samama, J.-P., Wallén, L. and Brändén, C.I., J. Mol. Biol., 146 (1981) 561.
5. Al-Karadaghi, S., Cedergren-Zeppezauer, E.S., Petrantos, K., Hovmöller, S., Terry, H., Dauter, Z. and Wilson, K.S., Acta Cryst. D50 (1994) 793.
6. Schmidt, J., Chen, J., DeTraglia, M., Minkel, D. and McFarland, J.T., J. Am. Chem. Soc., 101 (1979) 3634.
7. Dunn, M.F., Dietrich, H., Macgibbon, A.K.H., Koerber, S.C. and Zeppezauer, M., Biochemistry, 21 (1982) 354.
8. Makinen, M.W., Maret, W. and Yim, M.B., Proc. Natl. Acad. Sci. USA, 80 (1983) 2584.
9. Makinen, M.W. and Wells, G.B., Met. Ions Biol. Syst., 22 (1987) 129.
10. Dutler, H. and Ambar, A. in The coordination chemistry of metalloenzymes Bertini I., Drago, R.S. and Luchinat, C. eds., Reidel, London, 1983, p. 135.
11. Dutler, H., Ambar, A. and Donatsch, J. in Zinc Enzymes Bertini, I., Luchinat, C., Maret, W. and Zeppezauer, M. eds., Birkhauser, Stuttgart, 1986, p. 471.
12. Merz, K.M., Hoffmann, R. and Dewar, M.J.S., J. Am. Chem. Soc., 111 (1989) 5636.
13. Maret, W. and Makinen, M.W. J. Biol. Chem., 266 (1991) 20636.
14. Dworschack, R.T. and Plapp, B. V., Biochemistry, 16 (1977) 2716.
15. Eklund, H., Jones, A. and Schneider, G., in Zinc Enzymes Bertini, I., Luchinat, C., Maret, W. and Zeppezauer, M. eds., Birkhauser, Stuttgart, 1986, p. 377.
16. Cedergren-Zeppezauer, E.S., in Zinc Enzymes Bertini, I., Luchinat, C., Maret, W. and Zeppezauer, M. eds., Birkhauser, Stuttgart, 1986, p. 393.
17. Eklund, H., Samama, J.-P. and Wallén, L., Biochemistry, 21 (1982) 4858.

18. Bertini, I., Gerber, M., Lanini, G., Maret, W., Rawer, S. and Zeppezauer, M., J. Am. Chem. Soc., 106 (1984) 1826.

19. Maret, W., Zeppezauer, M., Sanders-Loehr, J. and Loehr, T.M., Biochemistry, 22 (1983) 3202.

20. Maret, W., Shiemke, A.K., Wheeler, W.D., Loehr, T.M. and Sanders-Loehr, J., J. Am. Chem. Soc., 108 (1986) 6351.

21. Maret, W. and Zeppezauer, M., Biochemistry, 25 (1986) 1584.

22. Corwin, D.T., Fikar, R. and Koch, S.A., Inorg. Chem., 26 (1987) 3079.

23. Werth, M.T. , Tang, S.- F. , Formicka, G., Zeppezauer, M. and Johnson, M.K., Inorg. Chem., 34 (1995) 218.

24. Hemmingsen, L., Bauer, R., Bjerrum, M.J., Zeppezauer, M., Adolph, H.W., Formicka, G. and Cedergren-Zeppezauer, E.S., Biochemistry, 34 (1995) 7145.

25. Boiwe, T. and Brändén, C.I., Eur. J. Biochem., 77 (1977) 173.

26. Sartorius, C., Dunn, M.F. and Zeppezauer, M., Eur. J. Biochem., 177 (1988) 493.

27. Andersson, I., Maret, W. and Zeppezauer, M., Brown, R. D., Koenig, S. H., Biochemistry, 20 (1981) 3424.

28. Andersson, I., Bauer, R. and Demeter, I., Inorg. Chim. Acta, 67 (1982) 53.

29. Maret, W., Zeppezauer, M., Desideri, A., Morpurgo, L. and Rotilio, G., FEBS Lett., 136 (1981) 72.

30. Makinen, M.W., Yim and M.B., Proc. Nat. Acad. Sci. USA, 78 (1981) 6221.

31. Bertini, I. and Luchinat, C., Met. Ions Biol. Syst., 15 (1983) 101.

32. Bauer, R., Adolph, H.W., Andersson, I., Danielsen, E., Formicka, G. and Zeppezauer, M., Eur. Biophys. J., 20 (1991) 215.

33. Formicka, G., Zeppezauer, M., Fey, F. and Hüttermann, J., FEBS Lett., 309 (1992) 92.

34. Maret, W. and Vallee, B.L., Method. Enzym., 226 (1993) 52.

35. Pocker, Y., Raymond, K.W. and Thompson, W.H., in Zinc Enzymes Bertini, I., Luchinat, C., Maret, W. and Zeppezauer, M. eds., Birkhauser, Stuttgart, 1986, p. 435.

36. Ryde, U., Int. J. Quant. Chem., 52 (1994) 1229.

37. Ryde, U., Proteins: Struct. Funct. Genet., 21 (1995) 40.

38. Singh, U.C. and Kollman, P.A., J. Comp. Chem.,7 (1986) 718.

39. Waszkowycz, B., Hillier, I.H., Gensmantel, N. and Payling, D.W., J. Chem. Soc. Perkin Trans. 2 (1990) 1259.

40. Waszkowycz, B., Hillier, I.H., Gensmantel, N. and Payling, D.W., J. Chem. Soc. Perkin Trans. 2 (1991) 225.

41. Waszkowycz, B., Hillier, I.H., Gensmantel, N. and Payling, D.W., J. Chem. Soc. Perkin Trans. 2 (1991) 1819.

42. Waszkowycz, B., Hillier, I.H., Gensmantel, N. and Payling, D.W., J. Chem. Soc. Perkin Trans. 2 (1991) 2025.

43. Böttcher, C.J.F., The theory of electric polarization Elsevier, Amsterdam 1973, p. 110,.

44. Halgren, T.A., J. Am. Chem. Soc. 114 (1992) 7827.

45. Ahlrichs, R., Bär, M., Häser, M., Horn, H. and Kölmel, C., Chem. Phys. Lett., 162 (1989) 165.

46. Teleman, O. and Jönsson, B., J. Comp. Chem., 7 (1986) 58.

47. Huzinaga, S., J. Chem. Phys., 42 (1965) 1293.

48. Wachters, A.J.H., J. Chem. Phys., 52 (1970) 1033.

49. Al-Karadaghi, S., Cedergren-Zeppezauer, E.S., Dauter, Z. and Wilson, K.S., Proteins: Struct. Funct. Genet., submitted.

50. Al-Karadaghi, S., Cedergren-Zeppezauer, E.S., Dauter, Z., Lamzin, V. and Wilson, K.S., manuscript in preparation.

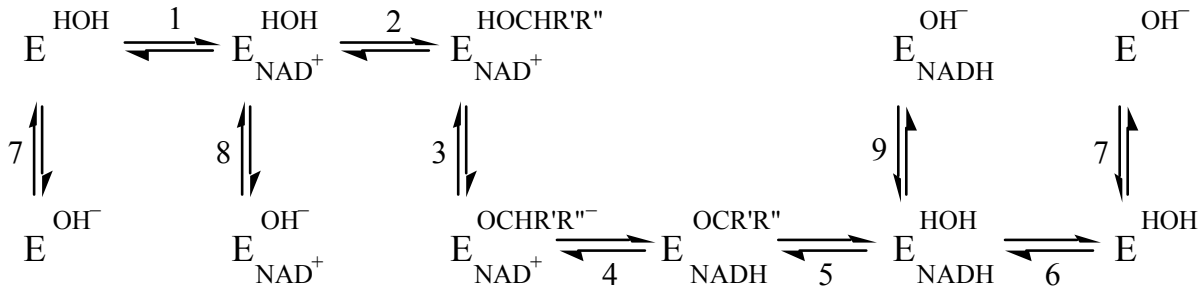
51. Bellido, M.N. and Rullmann, J.A.C., J. Comp. Chem. 10 (1989) 479.

52. Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A., J. Comp. Chem. 7 (1986) 230.

53. Liljas, A., Carlsson, M., Håkansson, K., Lindahl, M., Svensson, L A. and Wehnert, A., Phil. Trans. R. Soc. Lond., A, 340 (1992) 301.

54. Fraústo da Silvia, J.J.R. and Williams, R.J.P. The Biological Chemistry of the Elements, Clarendon, Oxford, 1991, pp. 182-184.

Scheme 1. Reaction mechanism proposed by Pettersson[1]. E denotes the enzyme.



Scheme 2. The flow of the combined quantum chemical and classical geometry optimisations. QC and CC denote quantum chemical and classical chemical, S1-S4 are system 1-4. The programs used are given in parenthesis [46,47].

0. Evaluate QC wave function of S1 including electrostatics of S2-S4 (DSCF)

Repeat

1. Evaluate the QC forces from S1 + electrostatics of S2-S4 onto S1 (GRAD)

2. Evaluate the CC forces from S2 onto S1, except the electrostatic interactions (MUMOD)

3. Add the CC and QC forces

4. Relax the geometry of S1 using these added forces (RELAX)

5. Change the coordinates of S1 in the CC representation

If S2 is to be relaxed then

6. Insert the Mulliken charges of S1 into the CC representation

7. Optimise S2 by CC energy minimisation keeping S1, S2 and S4 fixed (MUMOD)

8. Change the coordinates of S2 in the QC representation

9. Evaluate QC wave function and energy of S1 including electrostatics of S2-S4 (DSCF)

10. Calculate Mulliken charges of S1 (MOLOCH)

11. Evaluate CC potential energy of S2, except the electrostatics (MUMOD)

12. Add QC and CC energy

until change of energy and coordinates is below specified thresholds

Table 1. The performance of the model[a].

	QC	COMQUM fixed		COMQUM free		MM
system	bond	RMS	?bond	RMS	?bond	RMS
$MeOH+OH_2$	205.2	2.3	-3.8	2.1	-0.2	12.4
HOH+OHMe	201.3	3.9	0.4	2.7	3.3	33.6
MeOH+OHMe	202.6	2.6	-2.9	3.1	0.0	43.5
$EtOH+OH_2$	205.8	2.6	-1.1	7.1	-1.7	16.6
HOH+OHEt	200.9	2.4	2.7	13.7	-0.3	(145.8)
$MeOH+NH_3$	211.1	11.5	-1.7	4.6	-0.7	11.1
$H_2NH+OHMe$	234.4	2.6	-0.2	3.3	-0.6	41.9
HSH+OHMe	227.5	1.9	1.2	1.9	1.6	66.7
$HOH+SMe^-$	245.8	1.5	2.0	2.1	2.4	89.8
Average (with sign)		**3.5**	**1.8** (-0.4)	**4.5**	**1.2** (+0.4)	**39.4**

[a] Nine hydrogen bonded systems were geometry optimised quantum chemically (QC), molecular mechanically (MM) and by COMQUM with system 2 fixed or free to relax. In the COMQUM calculations the methyl and ethyl groups were treated classically and the other atoms quantum chemically. The table reports the root-mean-squared difference between the structures and the quantum chemical structure (RMS), the length of the hydrogen bond and the difference in this length (?bond) compared with the quantum chemical calculation. Unit: pm.

Table 2. Energies of structures optimised with COMQUM. B denotes $Zn(HS)_2$(imidazole)[a].

Complex	E_{tot} (H)		$?E_{QC1}$ (kJ/mole)		E_c (kJ/mole)	
$B(H_2O)$	-2877.907141	(673.1)	41.7	(5.2)	1062.2	(635.1)
B(H2O)[b]	-2877.878207	(665.0)	143.7	(24.3)	1125.4	(580.2)
$B(OH)^-$	-2877.315235	(624.4)	61.5	(1.4)	1134.9	(559.3)
$B(H_2O)_2$	-2953.849740	(626.1)	131.4	(18.4)	1119.5	(598.2)
$B(H_2O)+(H_2O)$	-2953.898436	(659.1)	63.4	(9.9)	1078.0	(620.0)
$B(OH)(H_2O)^-$	-2952.979733		172.5		1702.4	
$B(OH)^-+(H_2O)$	-2953.299701	(639.8)	69.9	(16.6)	1117.5	(580.5)

[a] A "+" in the formula indicates second sphere coordination. E_{tot} and E_c are described in Eqn 2 and 4, respectively. $?E_{QC1}$ is the difference in quantum chemical energy of system 1 without any classical charges at the geometry optimised in vacuum [35] and with COMQUM. The energies refer to computations where the enzyme is allowed to relax (except $B(OH)(H_2O)^-$ where no such minimum was found); values in parenthesis are the difference in energy between these calculations and the calculations with fixed enzyme (in kJ/mole).

[b] The water molecule occupies the alternative zinc site

Table 3. Geometries of structures optimised with COMQUM[a].

Complex	Protein	Distance to Zn (pm)					Angle subtended at Zn						
		N	S1	S2	O1	O2	S1-S2	S1-N	S2-N	S1-O	S2-O	N-O	O1-O2
Enzyme	A	214	224	235	219		130	113	104	106	102	94	
	B	205	222	227	216		129	113	108	105	102	93	
$B(H_2O)$	Vacuum	204	235	237	211		145	103	107	96	94	104	
	Fix	197	235	228	209		123	115	116	99	97	97	
	Free	199	238	233	212		124	115	116	99	97	96	
$B(H_2O)$[b]	Fix	206	237	230	223		112	124	119	96	119	78	
	Free	206	241	234	215		113	120	117	98	124	81	
$B(OH)^-$	Vacuum	210	246	247	187		108	132	96	112	132	95	
	Fix	205	249	241	188		111	107	106	121	105	106	
	Free	206	250	245	188		113	108	107	116	107	105	
$B(H_2O)2$	Vacuum	207	247	247	210	211	163	101	96	88	88	101	158
										89	88	101	
	Fix	207	246	235	212	220	108	127	125	88	90	91	164
										91	105	78	
	Free	204	246	238	223	231	110	127	123	83	89	90	158
										90	113	77	
$B(H_2O)+(H_2O)$	Vacuum	206	237	240	204	374	139	101	108	103	97	104	
	Fix	198	237	228	204	445	122	114	116	100	100	99	
	Free	200	241	234	206	431	125	114	115	98	101	98	
$B(OH)(H_2O)^-$	Fix	228	273	239	185	219	104	123	127	84	79	82	156
										120	98	80	
$B(OH)^-+(H_2O)$	Vacuum	210	243	247	190	384	108	101	102	124	116	102	
	Fix	199	248	236	190	437	113	110	111	110	103	110	
	Free	204	249	244	191	387	114	110	108	116	105	103	

[a] B denotes $Zn(HS)_2$(imidazole). A "+" in the formula indicates second sphere coordination. S1 and S2 in calculations including the enzyme represent the thiolates of Cys46 and Cys174, respectively. For the five-coordinate structures, O1 is the ligand in the substrate site and O2 in the one in the alternative site. For the enzyme, A and B refer to the two subunits.

[b] The water molecule occupies the alternative zinc site

Legends to the Figures

Figure 1. The junction between the system 1 and system 2, illustrated by a cysteine residue bound to the catalytic zinc ion in alcohol dehydrogenase. The H junction atom in system 1 corresponds to the CB atom in system 2 by Eqn. 1. To the quantum chemical energy and forces of system 1 is added molecular mechanical energy and forces due to the CB-HB1, CB-HB2, and CB-CA bonds, the SG-CB-X, CB-CA-X, and HBx-CB-X angles, the Zn-SG-CB-X, X-CB-CA-X, and CB-CA-X-X dihedral angles, and also the van der Waals interaction of any atom in system 1 (plus the CB atom) with any atom in system 2. Furthermore, the difference in molecular mechanical energy and forces of the Zn-SG-Y angle and the three Y-SG-Zn-X dihedrals with Y=CB and with Y=H is added to quantum chemical energy and forces, in order to correct the use of a H junction atom in the quantum chemical calculations (the SG-Y bond term cancels out due to Eqn. 1).

Figure 2. Comparison of the geometry of $EtOH+OH_2$, optimised quantum chemically (white), molecular mechanically (black) and with COMQUM with system 2 fixed (light grey) or free (dark grey). For details in the computations see the legend to Table 1.

Figure 3. Stereo views of the optimised structure of the active site of alcohol dehydrogenase with different ligands coordinated to the zinc ion. The amino acids of system 2 are shown. The enzyme was allowed to relax, except in f. The ligands are: a. one water molecule in the substrate site; b. one hydroxide ion in the substrate site; c. one water molecule in the substrate site and one in the second coordination sphere; d. one water molecule in the substrate site and one in the alternative site; e. one hydroxide ion in the substrate site and one water molecule in the second coordination sphere; f. one water molecule in the substrate site and one hydroxide ion in the alternative site.